Postel

Network Working Group
Request for Comments 1074

J. Rekhter
T.J. Watson Research Center
IBM Corporation
October 1988

The NSFNET Backbone SPF based Interior Gateway Protocol

Status of this Memo

This memo is an implementation description of the standard ANSI IS-IS and ISO ES-IS routing protocols within the NSFNET backbone network. Distribution of this memo is unlimited.

Acknowledgements

I would like to express my thanks to Hans-Werner Braun (MERIT) for his contribution to this document.

1. Overview

This document provides an overview of the NSFNET Backbone routing with specific emphasis on the intra-backbone routing.

By the end of 1987, the American National Standardization Institute (ANSI) forwarded a specification for an Intermediate System to Intermediate System routing protocol to the International Standardization Organizations (ISO) for the adaptation as an international standard. This ANSI IS-IS protocol is used as the interior gateway protocol (IGP) of the NSFNET backbone. Documented here is an implementation description which also includes further definitions that were necessary for the integration into an Internet Protocol (IP) environment. Therefore, it should be viewed as a continuation of the specifications of the ANSI IS-IS protocol [1] and the ISO standard End System to Intermediate System (ES-IS) protocol [2]. While the ANSI IS-IS protocol suffices as an IGP, additional methods are used to orchestrate routing between the backbone and the attached mid-level networks; most notably the Exterior Gateway Protocol (EGP). Further information about the overall NSFNET routing as well as some future aspects can be found in [3], [4], [5] and [6].

2. A brief overview of the NSFNET backbone

The NSFNET backbone is a wide area network which currently connects thirteen sites within the continental United States. All connections are permanent point-to-point links at T1 speed (1.544Mbps). These T1 links may contain multiple logical links at sub-T1 and up to the full T1 speed. The result is a hybrid circuit/packet switching network able to contain a connectivity-richer logical topology than the

underlying physical topology would allow by itself. Each site has a Nodal Switching Subsystem (NSS) which is responsible for packet switching. Each NSS is a RISC technology based multiprocessor system using IBM RT/PC processors which operate a modified version of a 4.3BSD kernel. For the purpose of routing, each NSS is considered as a single entity which has connections to both other NSS (via the logical network infrastructure) and to regional networks (via local area network attachments; typically an Ethernet).

The routing protocol which is used for the inter-NSS routing within the NSFNET backbone is an adaptation of the ANSI IS-IS routing protocol [1]. The routing protocol which is used between the backbone and the attached mid-level networks is the Exterior Gateway Protocol (EGP) [3]. The information exchange between the backbone and its connected EGP peers is subject to policy based routing restrictions which are maintained in the Policy Based Routing Database [4,5].

3. An overview of the ANSI IS-IS routing document

The ANSI IS-IS routing protocol specifies a two level hierarchical routing where Level 1 routing deals with routing within an area, while Level 2 routing deals with routing between different areas.

This routing protocol belongs to a class of so called "Link State" protocols where each node maintains a complete topology of the whole network. The route computation is based on a modified version of Dijkstra's Shortest Path First (SPF) algorithm.

Both Level 1 and Level 2 routing use two types of Protocol Data Units (PDU):

> The Level 1 Router Link PDU lists IS neighbors. The Level 1 End System PDU lists ES neighbors.
>
> The Level 2 Router Link PDU lists neighbor Level 2 routes. The Level 2 End System PDU lists address prefixes for systems in other Routing Domains.

The ANSI IS-IS document separates subnetwork independent functions from the subnetwork dependent functions. Subnetwork independent functions include dissemination of Router Link and End System Link PDU's and the Routing Algorithm. The subnetwork dependent functions cover different types of subnets such as X.25, permanent point-to-point links and LANs.

The IS-IS Protocol is designed to interoperate with the End System to Intermediate System (ES-IS) routing exchange protocol [2]. The ES-IS

protocol is used to determine connectivity and network layer addresses. This information is used to construct the Router Link PDUs.

4. How the ANSI IS-IS protocol is adapted for the NSFNET backbone routing

The NSFNET backbone implements a subset of the ANSI IS-IS protocol. With respect to subnetwork independent functions, it only supports Level 2 routing. With respect to subnetwork dependent functions, it only supports general topology subnetworks with permanent point-to-point links. Since the ANSI IS-IS protocol is designed for ISO Network Service Access Point (NSAP) addresses, there is a need to encapsulate IP addresses into NSAP addresses.

For this, the Initial Domain Part (IDP) is unused. The Domain Specific Part (DSP) includes nine bytes which are partitioned as follows:

2 bytes - administrative domain

2 bytes - empty

4 bytes - IP address

1 byte - empty

In the ANSI IS-IS protocol, each router has its own identifier (ID) which is 6 bytes long. For the NSFNET implementation, the first 2 bytes of the ID are empty and the last four bytes include the IP address of a particular router.

The NSFNET backbone PDUs (both IS-IS and IS-ES) are transmitted as a protocol on top of IP, with "85" being the assigned protocol number for this purpose. The IS-IS PDUs are distinguished from the IS-ES PDUs by the Protocol Discriminator Field within the PDUs. The IP fragmentation/reassembly mechanism provides support for transmission of up to 64 kilobytes in a single IP packet. Within the backbone, it is highly unlikely that the size of IS-IS PDUs will exceed this limit. Therefore, no IS-IS fragmentation/reassembly is implemented for this environment. This is different from the ISO framework where the ISIS is located directly on top of the Data Link Layer.

For the purpose of the NSFNET Backbone routing, each Autonomous System (AS) is treated as a separate Administrative Domain (AD). The list of administrative domains (as obtained via EGP and filtered through the Policy Based Routing Database) which are connected directly to a particular NSS is distributed in the set of the

partitionAreaAddresses part of the Level 2 Router Links PDU. Each area address is 5 bytes long and consists of 3 empty bytes (IDP) followed by 2 bytes of the Administrative Domain.

The reachability information obtained from regional networks via EGP is distributed within the backbone by End System PDUs. In order to support multi-domain topologies, the ANSI IS-IS protocol allows for a set of Address Prefixes to be entered by the System Management at the boundary IS. In the NSFNET Backbone, these Address Prefixes are obtained via the Exterior Gateway Protocol. For each network listed in EGP NR packets which is received from an EGP peer, the network and administrative domain number of the EGP peer are encapsulated into NSAP addresses (as described above). A complete NSAP address is used as an address prefix in the reachable address prefix neighbor part of the End System PDU. The cost field in the reachable address prefix neighbor part of the End System PDU is derived from the Policy Based Routing Database maintained in each NSS.

At each NSS, the reachability information obtained from other nodes (via their End System PDU's) is passed on to the mid-level network EGP peers, following the appropriate processing and filtering according to the Policy Based Routing Database.

The Network Entity Title (NET) (which is used in the IS-ES protocol) is eleven bytes long and is constructed by first encapsulating an IP address into a NSAP address, then taking the first 11 bytes of this address as a NET.

5. Current timer parameters

The following timer parameters are currently implemented:

Hello Interval (IS-ES Hello): 10 seconds

Hold Time (ES-IS protocol): 40 seconds

Other timer parameters for the IS-IS protocol are taken from the section 6.3.7 of [1].

6. References

[1] "Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol", ANSI X3S3.3/87-150R, 1987-10-29.

[2] "End System to Intermediate System Routing Exchange Protocol for use in conjunction with the Protocol for providing the Connectionless-Mode Network Service (ISO8473)", ISO JTC1/SC6/N4802R, 1988-03-26.

[3] Mills, D., "Exterior Gateway Formal Specification", RFC 904, University of Delaware, April 1984.

[4] Rekhter, J., "EGP and Policy Based Routing in the New NSFNET Backbone", IBM, March 1988.

[5] Braun, H-W., "The NSFNET Routing Architecture", Merit Computer Network, University of Michigan, April 1988.

[6] Braun, H-W., "NSFNET Inter Autonomous System Routing", Merit Computer Network, University of Michigan, September 1988.